DOCUMENT RESUME

ED 183 016 FL 011 029

TITLE	English Language Testing. General Information Series No. 20. Indochinese Refugee Education Guides.
INSTITUTION	Center for Applied Linguistics, Arlington, Va.: National Indochinese Clearinghouse and Technical Assistance Center, Arlington, Va.
SPONS AGENCY	Office of Refugee Affairs (DHEW), Washington, D.C.
PUB DATE	Sep 79
CONTRACT	ORA-600-78-0061
NOTE	35p.
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Achievement Tests; Bibliographies; Check Lists; Cloze Procedure; *English (Second Language); *Indochinese; Language Instruction; *Language Skills; *Language Tests; Listening Tests; *Refugees; Resource Materials; Second Language Learning; Speech Skills; Student Placement; Student Testing; *Test Selection; Writing Skills

ABSTRACT

Principles of test selection in English as a second language (ESL) are introduced to teachers of Indochinese refugees. No previous knowledge of ESL testing on the part of the teacher is assumed. A discussion of the characteristics of a good ESL test emphasizes the appropriateness of the test for non-native speakers, validity, reliability, and practicality. Specific types of tests are described, including: (1) discrete-point tests, exemplified by the Structure Tests-English Language (STEL) and Comprehensive English Language Test (CELT); (2) oral proficiency tests, such as the John Test and the Ilyin Oral Interview; and (3) Cloze tests as measures of readability and language ability. A guide to developing a strategy for language testing explains procedures for placement, progress, and final achievement assessment. Finally, a guide to classroom testing outlines actual procedures for administering listening, reading, speaking, and writing tests. A checklist of principles that should be observed in classroom testing is included. A bibliography of tests and teacher resources is appended. (JB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

National Indochinese Clearinghouse • Center for Applied Linguistics
1611 North Kent Street, Arlington, Virginia 22209

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Center for Applied Linguistics
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Indochinese Refugee Education Guides

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED183016

General Information Series #20: English Language Testing

Page

I. Introduction..... 1

II. Characteristics of a Good ESL Test..... 5

III. Types of ESL Tests..... 8

A. Discrete point tests..... 8
B. Testing Oral Communication Skills..... 11
C. Cloze Tests..... 14

IV. A Strategy for Language Testing..... 16

A. Entry/placement in the program..... 17
B. Assessment of progress and achievement..... 19
C. Leaving the program..... 20

V. Classroom Testing..... 21

A. Listening..... 22
B. Reading..... 24
C. Speaking/writing..... 25

VI. Summary..... 29

VII. Bibliography..... 20

I. Introduction

Language tests play an important role, in fact several important roles, in any program to teach English as a Second Language (ESL) to Indochinese refugees. The different types of tests often used are placement tests, achievement tests, and diagnostic tests.

Placement Tests. When a refugee enters an ESL program, it is important to place the refugee in a class at a level that will provide an efficient and succesful learning experience. Placement tests are used to help determine the appropriate class. (As we will see, other elements come into play as well, including the refugee's previous language training and educational experience,

FL011029

as well as the refugee's goals in studying English. It may be more appropriate to set up a class of students at somewhat different levels of proficiency, if they are all studying English for a more or less common purpose: to enter a vocational program, for example, as opposed to simply developing general English survival skills.) Placement tests may also give diagnostic information, both in a general and in a quite specific sense. If a battery of tests is used, it will be possible to assess a refugee's ability in various language skill areas. The refugee may have some ability to speak English and to communicate in English to some degree, but may at the same time have a much lower level of reading ability in English. Many refugees, in fact, are not literate in their own language, and, therefore, are not able to read English either. This sort of diagnostic information is important when deciding where to place the incoming language student. Many programs have found that, at least initially, it is wise to place illiterate ESL students in a class together, since, even if they have some oral fluency in English, they tend to have somewhat different instructional needs.

Achievement Tests. Tests and quizzes are part of the routine of the language classroom. For one thing, they let the teacher (and the student) know whether the student has mastered the material which has been taught during the class. Teachers often use more or less formal measures to assess their students. A short vocabulary test, for example, based on the words introduced in the language lesson, will tell the teacher whether the students have mastered that part of the curriculum. A teacher may put together a brief quiz on the structures introduced in a lesson to make sure that the students have control of those structures before moving on to another lesson. Language teachers frequently test their students' ability to tell the difference between two closely related English sounds, or between two sound patterns that indicate different meanings. This is not only language testing, but it is also a means of language teaching as well. Tests of this sort, which are familiar to us all, are the achievement tests that are part of any language teacher's repertoire. They are checks on how the students are doing, and they indicate whether the teacher has been successful in conveying the material to the class. Because they don't have to be very formal or rigorous to tell the teacher what is important to know, teachers can construct them with relative ease, and make them reflect the particular material that is being taught in a given class. Moreover, many ESL texts and series include sets of

achievement tests, keyed to the text and curriculum, as part of the resources available to the teacher.

Standardized achievement tests (often called proficiency tests) have another use. They are used to assess whether a student is prepared to enter other kinds of training or education programs, at least in terms of English language ability. Is the student ready, for example, to enter a college or university program intended for native speakers of English? There is a well-known comprehensive language test that was developed to measure exactly that: the Test of English as a Foreign Language (the TOEFL Exam, which is pronounced toe-full, with the accent on the first syllable). It is academic in its focus and vocabulary, and it is closely correlated to the actual performance of foreign students in American colleges and universities. The language tasks of a college student are fairly well defined, and the TOEFL Exam is used to predict whether a student has sufficient English proficiency to perform them.

However, many ESL programs for Indochina refugees are concerned, not with preparing students to enter colleges or universities, but rather to enter the job market or a vocational training program. What level of English ability does a student need in order to cope with a particular vocational training program, or a particular job situation? This is very difficult to determine, because it depends not only on the nature of the job, or the type of job training, but also on the determination, the previous experience, the resourcefulness of the refugee, and on many other individual factors.

If it is not possible to assess a refugee's language ability in a vocational context with the kind of precision that the TOEFL Exam provides for the college-bound student, it is possible to establish the refugee's general level of language proficiency. A number of standardized tests are available that measure general English proficiency, including oral English fluency. Although these tests are not keyed specifically to the work or vocational training context, they are certainly more appropriate as a measure of a refugee's proficiency (and therefore chances for success), than the tests intended for native speakers of English that are often used as entrance exams for vocational programs. Reading tests, designed for native English speakers, have little relevance in assessing the ability of an ESL student, whose language facility in English is still growing.

Diagnostic Tests. More sophisticated placement tests or test batteries and most standardized achievement tests can yield very specific diagnostic information about the refugee's command of English. They can indicate what English structures the refugee has mastered, and what areas of the language are likely to be particularly troublesome. Such information can help a language teacher plan the English language curriculum in order to give special attention to areas of particular difficulty. It is rare that an ESL test is written expressly for diagnostic purposes. For the very beginning student, a contrastive analysis of the sound system and grammatical system of his language with English can pinpoint those phonological and structural points with which the student will have the most difficulty.

Language tests, then, are used in these ways in ESL programs: as measures of achievement, or progress in mastering the language curriculum; as means for the placement of students at appropriate levels in the language program, and in some cases, for diagnosing particular language programs; and for establishing English language proficiency according to a fairly objective scale.

In this Guide, we will look at some different types of English language tests. For the most part we will be concerned with tests are are commercially available, and we will examine several closely. We will not go into the question of how to construct ESL tests, but we will include classroom type exercises that can be used for achievement purposes. Teachers who are interested in developing skills in test construction and test administration should look to one of the handbooks of resource manuals devoted entirely to making tests and validating them, such as David Harris' Testing English as a Second Language (New York: McGraw-Hill, 1969) or Rebecca Valette's Modern Language Testing, 2nd edition (New York: Harcourt Brace Jovanovich, 1977). Developing a test of general English proficiency that is a reliable and valid guide to making important decisions about a student's future is a complex and demanding task, and it requires not only specialized skills, but also long-term commitment of time and other resources. Many ESL teachers have found that published tests meet most of the needs of their assessment program, and they have concentrated their energies in developing achievement tests keyed to the particular curriculum they have used, or tests for certain situations where no published tests are appropriate, such as placement tests for illiterate refugees with low oral English skills.

We will examine closely some ESL tests that have proven particularly effective in measuring English language proficiency for Indochinese refugees, and we will show how they are used. We will also look at a general strategy for language testing in an ESL program, in light of the various functions and goals of language tests. Finally, the Guide will provide an annotated bibliography of ESL tests, as well as other resources for ESL teachers interested and concerned with questions related to using published tests and developing new tests for specific purposes.

II. Characteristics of a Good ESL Test

What should you look for when you are considering an ESL test? First of all, you should be sure that the test was designed for ESL students. That may seem like an obvious point, but people unfamiliar with language teaching often assume that <u>any</u> test of English is meant for ESL students. A reading test, for example, designed for English-speaking Americans is based on certain assumptions which make it inappropriate as a measure of the ability of a person whose native language is not English. It is based on the assumption that whoever takes the test has mastered all the structures of English that are used in ordinary conversation, and that he has a commonly shared background and general experience. But an ESL student may still be learning English structures, and the student's cultural experience may be quite different, so different that it interferes with the student's understanding of the material in the test. Here is an example taken from a reading test for native speakers of English:

> Corn and tomatoes were new to European tables when introduced by those returning from the New World. These vegetables came from
>
> America Europe China India

The ESL student is likely to encounter difficulties here not anticipated by the test makers. Several structural elements (e.g., "new to European tables" and "introduced by those returning from the New World") are typically considered high intermediate to advanced in ESL curricula, and they would likely be unfamiliar to many ESL students, even though they are assumed by the test makers to be part of the repertoire of the persons taking the test. Moreover, the cultural

references create difficulties for an ESL student from Asia. In short, this test item and the reading test it is drawn from wouldn't be a fair measure of the ESL student's ability to read English.

This is an example of one way a test can fail: it doesn't test what it is intended to test. It is not a valid test of ESL reading proficiency. Validity is an important aspect of any language test. Validity relates to these questions: What does the test measure? How well does it measure? For specialists concerned with constructing a test, these questions have very complex implications relating to defining the factors that go into making up a particular language skill, and to developing ways of testing those factors.

For the language teacher, the question of a test's validity has a somewhat different focus. Does the test measure a skill that is relevant for the students in this particular program? What does it tell the teacher about the student's developing mastery of English? Finally, how well does the test relate to other, outside criteria, such as success in a vocational training program or effectiveness on a job?

A well-constructed ESL test that is valid in one application may be inappropriate in another. Look at these examples. Suppose that a refugee with considerable education and some English skills enters an ESL program. The refugee wants an increased level of English proficiency in order to enter a vocational training program that requires good listening comprehension and good reading ability. The question is: at what level should the refugee be placed? One important aspect of placement would include diagnosing what structures the refugee has already acquired, and since the refugee was literate, a standardized structure test, related to the curriculum of the program, would be a valid test since it would indicate where to place the refugee in terms of the curriculum levels in the program.

However, suppose that the student, after spending some time in the program, was anxious to enter the training program that he had been preparing for. A structure test, which had been appropriate for placement into the ESL program, would not give a full and complete picture of the student's ability to cope with the language demands of the vocational training situation. A test of oral fluency including listening comprehension, together with some assessment of the student's ability to read the materials in the vocational program, would be a more valid test at this point.

A good language test should be reliable as well, that it, it should measure consistently and yield dependable results. The test should be capable, simply, of discriminating between students of high and lower levels of language ability. And the test should be stable, in the sense that a student who takes the test on one day should get about the same score on an equivalent form of the test taken on another day.

Constructing a test that has a high degree of reliability can be a real challenge for test specialists, especially when they are concerned with developing a test of language elements or general proficiency that will be given to large numbers of students. Even the classroom language teacher, dealing with a small group of students with familiar capabilities, has to be concerned with using reliable and consistent tests and quizzes. This means that all the students need to be measured according to a common set of standards. Careful teachers use formal tests to check the evaluation they make of their students in informal ways.

A language test is said to be reliable when it measures with precision. Now that is not the same thing as saying that the test actually measures what it was intended to measure (which has to do with the test's validity), but merely that it is consistent and dependable. A structure test would be called reliable if it consistently indicated which students had mastered a given set of English structures, and which ones had not.

A final essential characteristic of a good language test is that it should be practical to use and to score. Some test instruments that are extremely consistent in the ways they measure, and that are quite valid measures of true communicative competence in English, are nevertheless impractical for most ESL programs. The "scored interview", which is described below, is an example of an extremely powerful test that is simply not practical for most language programs. It is too expensive: it requires more than a man-hour to give and score the interview to a single student, and, moreover, it has to be scored by highly trained (and expensive) test specialists. Somewhat less reliable tests, which are nevertheless designed to focus on the same sort of language competence, have to be chosen instead, because they are far more practical.

So these characteristics of good language tests have to be taken into consideration: their validity (whether they measure what they are supposed to measure), their reliability (whether they measure consistently), and their practicality (whether they can be given economically). The choice of what tests to use involves some compromise among these points. The types of ESL tests we will consider, and describe in the Bibliography, generally meet these criteria, if they are used in appropriate contexts.

III. Types of ESL Tests

There are two basic approaches to testing language proficiency. One approach assumes that language proficiency, like language itself, is made up of identifiable, separate, discrete elements, and these elements may be measured individually. Tests that focus on the individual elements and measure them more or less in isolation from other elements are called discrete-point tests. Another approach is based on a different assumption: that the ability to communicate is a global ability, and that enumerating and measuring the control of individual elements does not necessarily add up to an assessment of true communicative competence.

Tests based on both approaches are very commonly used in ESL programs. Discrete-point tests have some advantages in terms of ease of administration and the kinds of diagnostic information they provide. More global measures of communicative ability can offer more insight into how a language student is likely to actually perform, that is, actually use the language to get the message across. For these reasons, many ESL programs have adopted a test strategy that employs both types of testing in order to get a more comprehensive picture of an ESL student's proficiency in English.

A. Discrete-Point Tests

Discrete-point tests measure a student's control of specific elements of a language. Essentially, they test one element at a time. One test may focus on aspects of English structure, another on vocabulary items, while others may measure the student's ability to discriminate among English sounds, or to produce them. Commonly, discrete-point tests are given using a multiple-

choice (or fill-in-the-blank) format, but other forms may be used as well. For example, a writing test or even a speech sample may be scored in a discrete-point manner. That is, only a particular aspect of the writing or speech sample would be considered: structural elements, for example.

Most published language tests are considered to be primarily discrete-point tests: they measure specific elements of the language. For this reason, they are powerful diagnostic instruments, for they offer very precise information about what areas, or specific elements of the language the student has mastered. They are used for placement at various levels in an ESL program, for diagnostic purposes, and for assessment of general proficiency (on the assumption that control of the discrete elements of a language eventually add up to general communicative competence).

A number of discrete-point language tests have been used for a variety of purposes in ESL programs for Indochinese. The STEL Test (Structure Tests-English Language) developed by Jeanette Best and Donna Ilyin is a good example of a discrete-point approach to language testing. The STEL is available for three levels--Beginning, Intermediate, and Advanced--and there are two forms of the test for each level (a useful feature, if one wants to give a pre-test and a post-test).

Each form of the STEL contains 50 items that are designed to test the student's ability to identify correct English structures. Each item consists of three sentences which are identical except for one underlined element. The student marks the one that is grammatically correct. This example is taken from the introductory material:

A He <u>is</u> a student.
B He <u>am</u> a student.
C. He <u>are</u> a student.

Students mark the correct answer on a separate answer sheet.

The STEL is based on a rigorous analysis of the structures of English, and a sequence of increasingly complex structures underlies the organization of the different levels of the tests. However, the indications of "Beginning, Intermediate, and Advanced" are by no means absolute terms. Students who might be considered advanced in one program could be the intermediate students in

another, perhaps larger program. The levels of the STEL are carefully described in the manual which accompanies the tests, and scores on the STEL are correlated with other published tests, but the levels should be carefully considered with your particular program in mind. After gaining some experience using the STEL, or other similar discrete-point tests, what the levels indicate in terms of the needs of your program will become clearer.

Discrete-point tests are by no means limited to the assessment of a student's knowledge of English structures. They may be used to measure other elements of language proficiency, and other language skills. A test battery, such as the CELT (Comprehensive English Language Test) is made up of three individual tests: a Listening Test, a Structure Test, and a Vocabulary Test. Each test focuses on a different aspect of language ability. The Listening Test measures a student's ability "to comprehend short statements, questions, and dialogues" spoken by native speakers of English. The Structure Test assesses one's ability "to manipulate the grammatical structures occurring in spoken English", and the Vocabulary Test measures the student's knowledge of "lexical items which occur in advanced English reading". The CELT materials were developed for, and have been normed against ESL students in academic situations, high schools and colleges. For this reason, the focus and tone of the tests are academic in nature. The level of the CELT is also somewhat higher than the lower forms of the STEL.

Most of the items in the CELT are typical of the usual form of discrete-point multiple-choice tests. Each item presents a "stem", a sentence that offers a context, and is followed by a multiple choice item. This example is taken from the introduction to the Structure Test:

"How old is George?"

"He's two years younger _____ his brother Paul."

(A) that
(B) of
(C) as
(D) than

Both the Vocabulary and the Listening Tests are organized in essentially the same way. The Listening Test may be given either with the test administrator reading from a script, or using a recorded tape that accompanies the test.

The CELT, like the STEL, and other well-made discrete point tests, such as the Michigan battery, have a number of attractive features. They are easy to give and to score. They may be administered economically, since they can be given to large groups at one time. They yield rather precise information, and, because they have been developed over a long period of time and with great care, scores on the tests are generally accepted as reliable and valid in well-defined contexts. However, these tests were developed, as we have said, with an academic focus, and they are less relevant to other contexts. They have demonstrated a high level of accuracy in predicting success in school work in college and university settings. But they may be less effective at identifying and measuring more global language skills that are involved in communicating effectively in English in less formal settings. Other, more global measures of communicative performance may need to be used to add more information about an ESL student's proficiency.

B. Testing Oral Communication Skills

For most students in an ESL program, the main goal of language instruction is to develop the ability to speak with and to be understood by native speakers of English. The accurate assessment of a student's ability to use the language for oral communication has been one of the most difficult problems for the test maker and the language teacher. The traditional means of testing for communicative ability was the use of discrete tests that assessed various elements that make up language: vocabulary, structure, listening comprehension, and the production of English sounds and sound patterns. It was assumed that these elements "added up" to general language proficiency.

Some tests were used to assess speaking and communication skills directly. The U.S. State Department, for example, has used for many years a kind of language test known as the "scored interview". Basically, it works like this. A student carries on a conversation about general topics and also about very specific areas of professional and personal interest with two or more native speakers of the language, who are highly trained in assessing aspects of language proficiency. The conversation may last for half an hour or more. On the basis of this, it is possible to measure with great precision the student's command of the language. The problems with "scored interviews" is that they are simply

too time-consuming to be practical for most ESL programs, and they require a high degree of training on the part of the scorers.

Fairly recently, other tests have been developed to assess oral communication skills that are much more practical and easy to administer. They are called "structured oral interviews", and they are designed to elicit from the language student a range of increasingly difficult structures, vocabulary, and uses of language in a setting that is more or less "natural". Because they can be scored objectively, they don't require highly trained language test specialists to do the evaluation.

There are two structured oral interviews that are widely used in ESL programs for Indochinese students: a brief, rather simple oral test, The John Test (after the name of the character, John, who is used as a focus for questions in the test), and the Ilyin Oral Interview, which is a longer, and much more sophisticated test instrument.

The John Test was specifically designed to test oral fluency of adult ESL students. It is a short oral interview, based on a set of pictures that illustrate activities of a character named "John" during a typical day. The interview is divided into two parts. In the first part, after pleasantries have been exchanged, and the purpose of the test explained, the interviewer asks the student a number of questions, eleven or twelve in all, about "John" and his activities which are shown in the set of seven pictures. The student's responses are graded as correct in fact and grammar, or correct in fact, but with some grammatical error, or factually incorrect. In order to make the interview as natural as possible, some test administrators accept short form answers, but others require students to answer using complete sentences. (The latter is less natural, but it does provide more information about the student's command of structures.) Obviously, only a very few structures are sampled in the John Test, but the test does sample a range of structures of increasing difficulty.

In the second-part of the test, the student is asked to relate a narrative based on the pictures. The connected discourse that the student produces is rated according to "fluency" and to control of English structures, on a scale of 0 to 14. Although this seems a good deal more subjective than the scoring

for the first part, the test makers have reported a high degree of scoring consistency among test administrators who are familiar with and experienced in giving the test.

The John Test provides a direct assessment of oral fluency skills, with enough precision so that the performance of students can be compared with considerable objectivity. The test is especially good at identifying students who are able to communicate fairly effectively, even though their command of English structures is pretty deficient. Teachers who deal with Indochinese students in ESL programs that emphasize oral communication skills have found that the John Test is quite useful for placing students at appropriate levels. One additional advantage of the John Test is that it does not seem to require much training to learn how to administer. It can be given, the instructions say, by any native speaker of English with a minimal amount of training and practice. Counselors, para-professionals, and volunteers can be trained to give the test, and this is an important consideration when large numbers of students have to be tested. (The John Test, like other oral interviews, is designed to be given individually.) The test makers suggest, however, that a trained and experienced ESL teacher is likely to get more information about a student when giving the test, simply because of the teacher's greater understanding of what is involved in language proficiency.

The Ilyin Oral Interview is, as we have said, a much more sophisticated test of oral proficiency. Like the John Test, it is based on a series of questions relating to a set of pictures. But is not only includes many more questions (30 in the short form, and 50 in the full form), but the questions are based on a much more rigorous-sampling of the structures of English. For this reason, the Ilyin Test can be correlated closely with other types of tests, such as the Michigan Structure Test and the CELT Exams. The Ilyin Test also provides a good deal of specific diagnostic information. The results of the test give a clear picture of what structures a student has acquired and are part of the student's active oral competence in English.

However, because of the greater sophistication of the test, many teachers have found that it is considerably more difficult to learn to administer the Ilyin Test with confidence. It takes many hours of practice to learn to give

the test smoothly. The Ilyin Test is also much more time-consuming to administer, since it takes up to half an hour to give and to score.

These two tests have been very successful instruments for assessing directly the oral fluency of ESL students. It should be said that neither test ought to be given to a student who lacks minimal proficiency in English. Unless the student can deal with the initial questions on the test, the exam should be terminated, since otherwise it would become an exercise in frustration for both the student and the test administrator. (Both tests make provision for smoothly ending the test when the frustration level is reached.) Some ESL programs have developed modified versions of these tests (especially the John Test which is much less formal), which have been adapted to the particular needs of the program. The great advantage of "structured oral interviews" is that they provide direct information about the language skills that are most relevant for most teachers and students: proficiency in oral English.

C. Cloze Tests

Close tests are based on the nation that human beings tend to perceive things as wholes; if something is missing, people tend to fill in the gaps. A cloze test is simply a reading selection in which certain words have been deleted in a mechanical manner. Typically, every 5th word (or 6th or 7th word, etc.) is left out, and students are asked to fill in the deleted words. Total objectivity is observed when selecting the words to leave out, and there is no consideration of context or importance of the word omitted. Usually, the first sence or two is left complete, then every nth word is deleted from the remaining sentences. Here is an example taken from Valette (Modern Language Testing).

Fill in the missing words.

The great murder wave of the 1970's appears to have ebbed at last in big-city America. _____ reports from police departments _____ twelve selected cities show _____ in nine of them, _____ number of homicides dropped _____ -- and in some cases _____ -- last year. The drop _____ have halted for the _____ being a steady upward _____ in killing that reached _____ peak in 1974, the _____ lethal year since uniform _____ statistics have been kept _____ the United States.

(Correct responses: First; in; that; the; markedly; sharply; may; time; trend; a; most; crime; in.)

The test is extremely easy to score. The students are given a point for each correct response. One slightly different scoring system is used, in which a response is counted correct if it is grammatically acceptable and makes sense in the context of the passage. This second means of scoring is used more often with ESL students.

A cloze test is thought to be a measure of a student's global language ability, since it touches on points of structure, vocabulary, and comprehension in a general sense. Interestingly, the results of cloze tests seem to correlate with the results of listening comprehension tests. But mainly these tests, which are very easy to prepare and score, provide important information about a student's general proficiency in dealing with actual reading passages.

Cloze tests were developed originally to measure the readability of a written text. A teacher or researcher would select a passage from a book or other written material and delete words from that selection, using the mechanical cloze procedure. Then the selection would be given to a group of students, and from the results of this brief test it would be possible to determine whether the book as a whole was too difficult for the students to read. Only later was the cloze test procedure used to measure the reading ability of students.

In the context of an ESL program, cloze tests can be used for both purposes: to measure the readability of a text for ESL students at a given level, and to test individual students' general English language proficiency. Suppose that a teacher wants to use some additional reading materials in a class, but doesn't know whether the materials would be too difficult to be profitable. The teacher can solve this problem by using the cloze procedure with a selection from the materials, and give a brief cloze test to the whole class. It is usually estimated that if students on the average get less than 45% correct, then the material is likely to be too hard to be used as supplementary reading. If they score between 45% and 60%, then the reading would be suitable to be used with teacher supervision; and if they score higher than 60%, then the materials are a good source of free outside class reading.

This would be particularly useful for classes at higher levels, in which it is valuable to incorporate reading materials drawn from the kind of reading the students will have to cope with when they leave the ESL program. There are few really pertinent ESL reading materials for upper level students, particularly

in technical or vocational areas. However, some reading materials for native speakers of English are easy enough for more advanced ESL students. The cloze test procedure can help teachers identify such materials.

Cloze tests can also be extremely useful measures of whether a student can deal with the actual language demands of an education or training program. Short cloze tests based on the materials used in the training program and given to the student will indicate quickly whether the student can cope with language at that level. Using the cloze test procedure in this way, it is possible to develop a test instrument that is keyed to the language, including the vocabulary and typical methods of textbook organization, of any technical field.

It appears that cloze tests have a number of very promising applications in an ESL testing program. They are relatively easy to construct and to administer and score. They can be based on quite specific reading tasks, and they seem to provide teachers with a view and assessment of a student's general ability in English.

IV. A Strategy for Language Testing

Effective ESL programs incorporate a regular strategy of language testing into the design and daily workings of language instruction. Part of the purpose of testing is to keep tabs on the progress of the students, and to make sure that students are in appropriate classes for efficient learning. But testing also serves the goals of the language students as well. If there is a regular and predictable pattern of testing and assessment, students are given an added insight into their own progress, an insight that most students welcome.

Adopting a testing strategy implies identifying a variety of language tests and assessment procedures, and using them for a variety of purposes. Testing falls into three stages, and somewhat different approaches may be used for each stage. The first stage is when the student enters the program. The second concerns the time the student actually spends in the program developing language proficiency. And the third stage is when the student is preparing to leave the program for school, a job, or other training. Different testing strategies may be employed at each stage.

A. Entry/Placement in the Program

Suppose there is an ESL program of moderate size consisting of four or five classes at roughly three levels, intermediate and lower. Some of the classes have a vocation-related focus, and others are concerned with more general, survival English. A substantial number of students at the lowest level are not literate in their native language. In short, a fairly typical ESL program for Indochinese refugees. Ten or so new students enter the program at about the same time. What sort of test startegy should be used to place them in classes?

The goal of placement testing is to put each student in a class that will support effective and successful language learning. A number of different factors will affect this, and the placement testing should account for them. In this program there are different classes for survival English and for vocation related English, and this implies that each student's purpose for studying English needs to be taken into consideration. Moreover, the issue of literacy, which seems to be an important factor in predicting progress and the rate of achievement, must be considered as well. Finally, since the program has three different levels of language instruction, the proficiency of each student is a factor as well. It is important to notice that of the factors affecting placement, only one is specific to language proficiency.

A placement strategy is essentially a screening process. In screening for language proficiency, many programs have adopted a procedure which uses a number of language tests arranged in sequence. Initially, a determination is made whether the potential student has any English proficiency at all. A brief interview, in a non-threatening atmosphere, is arranged. At first, this may be a simple exchange of names and greetings. A student who demonstrates some command of basic English may then be given a short oral interview of the structured type, such as The John Test. In fact, The John Test begins with a set of social pleasantries, the purpose of which is both to put the student at ease, and to establish minimal English competence. If the student can procede in English, then the full test may be given. The results of the oral test should indicate the approximate level of the student in relation to this program.

But, as we said, other factors are also important for placement, the student's goals, for example. Where there is the possibility of using bilingual staff, entering students are interviewed in their own language in order to review their previous educational and work experience, and to get an indication of their reasons for entering an ESL program. A class in survival English would be a frustrating experience for a student who had well-defined goals in a vocational area.

An examination of the student's educational background and experience will give strong indication of the student's literacy, but some programs have also tested basic functional literacy in English and in the student's native language (assuming that it has a literate tradition) by simply asking the student to fill out simple forms in each language.

These screening procedures should provide enough information for placing a student appropriately in this small program. But it is wise to let the student also have a voice in the decision, and let the student move to another class if it seems a better arrangement after a time.

Some programs use the initial intake period as a time to establish a baseline estimate of a student's proficiency, a standard to measure the student's progress by. So additional tests are administered to entering students who have a sufficient level of language skills. The STEL is used widely for this purpose, as well as to gain additional placement information, if the program is large enough to require it. The Michigan Structure Test, and the CELT battery for more academically oriented programs, are also in wide use. Another placement test has also been effective for initial placement in many programs, even though it is keyed to a particular set of text materials, the Placement and Proficiency Test Package for Orientation in American English. Even program that don't use the OAE text have found the tests useful indications of levels of proficiency.

Entry and placement procedures, then, are used to get a comprehensive profile of the entering student, so that a program may be planned that is suitable and effective in meeting that individual's needs.

B. Assessment of Progress and Achievement

Teachers use a variety of means to assess students while they are studying in the program as well. Most teachers find they have to spend some time developing short tests and quizzes, because no standardized tests can fully capture the progress of students in an individual language class. The books by Harris and Valette treat specific questions of teacher-made tests and the test construction process, and in the following chapter we will give some possible test formats.

In addition to teacher-made tests, many teachers rely on the tests that accompany many ESL texts and series of materials. These tests are based on the vocabulary introduced in the materials, and they closely follow the sequence of structures on which the materials are based. ESL: A New Approach for the 21st Century (MODULEARN), to take one example, includes a test with each of the 40 lessons in its Beginning Level text. The tests include structure recognition, a writing test, and an oral segment. Every fifth lesson is accompanied by a test that reviews the previous five lessons. Tests may be ordered for other series also, including the widely-used English for Today from McGraw Hill (the Teacher's Manual for this series is extremely comprehensive, and it contains substantial information on how to write and administer tests related to the material in the texts), and IML's Orientation in American English.

As students move to somewhat higher levels, the more comprehensive placement tests can slso be used to assess achievement, when they are administered periodically. The STEL, which is available in two forms for each of three levels, has been used by some programs as a test of progress, though a fairly substantial amount of time must elapse before students begin to show a great deal of progress on any standardized test.

Whatever strategy of achievement testing, or combination of strategies, is employed, it should be remembered that regular, systematic testing is a good motivator for many students. It is crucial, however, to test the skills that have been taught, the skills that are given the highest priority in the instructional program. If the primary goal of the program is to develop communicative competence and oral fluency in English, then the testing program should reflect this goal. Students are very likely to study what they are going to be tested

on. If the class emphasizes oral communication, but the tests concentrate on the knowledge of grammar, then the tests are almost surely going to undermine the purposes of the instruction. The tests should support the instruction, not work against it.

C. Leaving the Program

One of the most difficult aspects of language testing is determining when a student is sufficiently proficient to leave the language program. Only a few language tests, such as the TOEFL exam, have a high degree of predictive validity for specific contexts, such as work in a college or university program. Once again, the assessment is based only partially on purely linguistic grounds. Other factors, such as the student's motivation, and preparation in other areas, are sure to have a powerful effect on each individual's degree of success.

Nevertheless, some of the standardized tests we have considered do offer a detailed and comprehensive view of a language student's general level of proficiency. The use of global measures, such as an extensive oral interview, like the Ilyin test, seems particularly important at this stage. It is vital to know how well the student can actually use English to communicate in a natural setting. The structured oral interview, although not exactly an ordinary language occasion, still simulates the actual demands of language use in the real world.

We have also suggested that cloze tests, based on materials drawn from the job or the training environment, could be developed quickly and rather easily, and they could be highly individualized, since they treat the specific language requirements that the student will actually face.

Finally, teacher assessment will come into play, based on the teacher's familiarity with the student's work, abilities, and progress through the language program. Careful teachers, as we have said, will use the resources of a wide variety of test procedures to ensure that their judgment is an informed one.

V. Classroom Testing

As mentioned before, classroom testing of achievement is desirable from both the students and the teachers point of view. If the basic text material being used is not accompanied by tests specific to that material, the teacher will have to devise testing situations. And even if tests do accompany text materials, the teacher should be able to devise alternate testing strategies. The following is a checklist of principles that should be observed in classroom testing.

(1) Test what has been taught.

(2) Test the objectives of the course.

(3) Tell students specifically what material is to be covered on the exam.

(4) Familiarize students with test format before giving examination.

(5) Check to see if directions are clear.

(6) Test one item at a time whenever possible.

(7) Try to test in context.

(8) Test all language skills: reading, listening, speaking, and writing.

(9) Make each test a representative sample of material taught.

(10) Weigh exam in accordance with the stated objectives of course.

(11) If possible, consider ease of correction as well as administration.

(12) View exams as a learning experience for both the teacher and the student. Help student identify his strengths and weaknesses. Moreover, provide the student with specific and supportive suggestions whenever possible.

What follows are some possible testing formats for assessing listening, speaking, reading and writing in the classroom. Subject matter and degree of difficulty will, of course, change according to the level of the class, but these item types can be used at any level for a check on student progress. Further types can be gotten from the Harris and Valette and Bartz books listed in the bibliography.

A. Listening

Type: Phonological Discrimination

Purpose: To test students' ability to recognize and compare:

a) sounds (minimal units of contrast)

b) intonation contours

c) stress

Example: Teacher reads contrastive or similar units and asks students to state whether the elements they heard were the same or different, question vs. statement, intonation emphatic vs. normal stress, etc. (This is much like the pronunciation exercise types used to teach pronunciation.)

Students may respond by writing the number 1 if one element were the same and 2 if they contrasted. Additional points, that is, a penalty factor should be subtracted for wrong answers to guard against indiscriminate guessing. For example, one might assign two points (+2) for each correct answer, and minus three points (-3) for each incorrect response.

Type: Appropriate Response

Purpose: To test students' ability to respond appropriately in an oral message.

Example: Teacher reads: What did you think of the soccer game?

Student reads:

a) It was the most boring game I ever saw.

b) I thought of the game.

c) It was the most boring game I ever went.

Type: Global Comprehension

Purpose: To test students' ability to hear a small segment of discourse and make global inferences as to where the conversation took place.

Example: Teacher reads: How much is the lettuce? Do they sell rice? Let's ask the manager. Where does this conversation take place?

Student reads:

a) It's a nice place.

b) In a supermarket.

c) In an airline office.

Type: Statement Rejoinder

Purpose: To check students' ability to respond with an appropriate rejoinder to an oral stimulus.

Example: Teacher reads: Would you mind if I took your plate now?

Student reads:

a) Yes, I am finished eating.

b) Yes, I haven't finished yet.

c) No, I am still eating.

d) No, I haven't finished yet.

Type: Completion

Purpose: To verify students' ability to complete logically an utterance presented orally.

Example: Teacher reads: I'm hungry.

Student reads:

a) Where is the bank?

b) Where can I get something to drink?

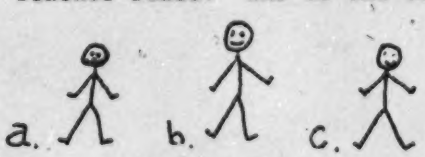
c) Where is the bathroom?

d) Where is the nearest restaurant?

Type: Comparisons

Purpose: To test students' ability to listen to an oral description and find one corresponding visual representation of the utterance given.

Example: Teacher reads: Who is the tallest?



Student reads: a) Paul b) Jim c) Fred

B. Reading

Type: Cloze.

Purpose: To check students' reading comprehension and his ability to supply missing forms when reading a passage.

Format: Teacher selects a short brief reading passage and deletes every fifth or seventh word.

Task: To supply missing words.

Type: Reverse semantic cloze/confused language.

Purpose: To evaluate students' ability to disregard irrelevant information. Also, this exercise can be used as a speed comprehension test.

Task: To cross out all irrelevant words (in a given time frame).

Example: The students enjoyed them the party very much. They stayed there while a long time. In fact they candy didn't leave until happy two o'clock in the sunset morning.

Type: Logical inferences

Purpose: To evaluate students' ability to make logical inferences based upon a reading passage.

He went to bed early because:

(a) He was tired.
(b) He was busy.
(c) He likes music.
(d) The movie was good.

Type: Completion

Purpose: To evaluate students' understanding of discrete grammatical or lexical items.

Examples:

1) Grammar

1. I would like _____

a) going a job
b) to get a job
c) job
d) would get a job

2) Vocabulary

1. I enjoyed the book very much
   a) the reading
   b) the movie
   c) the argument
   d) the talk

Variation:

1) He's my sister's husband. He's my _____.
2) Vinh is using an umbrella. It _____ raining.

Type: Same - Different

Purpose: To test students' ability to differentiate between grammatical or lexical forms.

Example:

Indicate whether the pairs of statements that follow are the same or different by writing S (same) or D (different) on the line provided.

1. ____ He's not old enough to drive.
   He's too young to drive.
2. ____ He's hardly working.
   He's working hard.
3. ____ He could have helped.
   He might have helped.
4. ____ He could not have said that.
   He might not have said that.
5. ____ He must not go now.
   He doesn't have to go now.

Note: A penalty factor should be built into the scoring of such items to discourage indiscriminate guessing.

C. Speaking/Writing

The testing of both speaking and writing skills presents different challanges for the teacher than the testing of listening and reading, largely receptive skills. In the assessment of productive skills, it becomes imperative that the standards for the evaluation of performance be clearly defined so as to minimize the somewhat subjective aspects inherent in the

rating of speaking and writing. Once operational definitions and guidelines have been established, a more objective evaluation is possible.

Speaking Tests

Type: Directed dialogues.

Purpose: To test students' ability to create natural conversations with a minimum of errors and a fair degree of fluency.

Format: Teacher reads or students read a brief incident from which they have to create a dialogue.

Situation: A young man, dressed in jeans, is being questioned by a clerk in a employment office.

Type: The telephone game

Purpose: To check students' understanding of roles and functions in the target culture. Moreover, this activity checks students' ability to ask questions.

Format: Teacher asks students to pretend to telephone the following places:

(1) police department
(2) employment office
(3) fire department
(4) restaurant
(5) school

Type: Directed discourse

Purpose: To evaluate students' ability to ask questions in English.

Format: (To student A):
Ask student B if he's ever eaten spaghetti.
Ask him if he liked it.
Ask him where he ate it.
Ask him what it tasted like.

Variation: The teacher knows an individual in the class who has done something "unusual" recently. He encourages other students to generate questions about his "achievement."

Examples:
Trip to Montana
Visit to the Grand Canyon, or New York City, etc.
etc.

Type: Games/Variation
Purpose: To check students' ability to formulate questions in English.
Format: By using the format of such games as:
"What's My Line?"
"Twenty Questions"
"I've Got a Secret"
The teacher can evaluate the students' questioning strategies.

Type: Interview
Purpose: To check on students' ability to generate questions and to expand appropriately upon information given.
Format: Student is given a blank application form (for credit or employment). He is then directed to ask questions in order to complete the application form.

Type: Role Plays
Purpose: To check students' ability to describe what he sees in clear and accurate English and to communicate that description to another person effectively.
Format: Teacher gives student A a picture which student A must describe to student B. Student B tries to draw what he hears. At the end of the task, student A and student B compare pictures.

Type: Giving Directions/Map Skills
Purpose: To check students' abilities to give clear and accurate directions.
Format: Teacher provides students with maps, or he may use a wall map/poster.
Task: Student is to describe how to get from Location A to Location B.

Type: Outlines
Purpose: To check students ability to expand dehydrated sentences into complete ones.

Format: Students are given an outline on which they must expand.

Type: Guided Speaking

Purpose: To check students' ability to communicate effectively in English -- to make descriptions, to report his feeling, etc.

Task: Have student describe a meal they had in a restaurant, a movie they saw, or an occasion in their life when they felt they were in danger.

Writing Tests

In most ESL programs for refugees, writing is only a small part of the course design. Below are some simple ways to test beginning writing.

Type: Dictation

Purpose: To determine if students can record in correct, grammatical English what they have heard aurally.

Format: Teacher should be consistent in his giving of dictation to insure comparability. One such procedure is as follows:

(1) Teacher reads entire selection at normal conversational speed.

(2) The passage is then divided into natural phrase sequences, with sufficient pauses given to allow students to write down phrases.

(3) Finally, the entire selection is re-read at normal conversational speed.

(4) Following final reading, the teacher should allow two or three minutes for students to review their papers and to make any revisions, if necessary.

Type: Sentence Builder

Purpose: To evaluate students' mastery of syntax by building complete sentences with dehydrated forms.

Directions: Combine the words, adding elements, if necessary, to make a complete sentence in English.

Example: TIM/SAN FRANCISCO/GO/LAST YEAR/VACATION

Answer: Tim went to San Francisco last year for his vacation.

Type: Controlled Composition

Purpose: To evaluate students' ability to transform questions into statements so as to write a coherent composition.

Example: Write a paragraph by changing the following questions into statements:

Did everything go wrong for Jack yesterday? Did he oversleep because he didn't hear the alarm clock? Did he get up quickly? Was he late for work? (etc.)

Type: Completion/Verbs

Purpose: To evaluate students' ability to supply correct verb forms in sentences.

Format: Complete the sentences with the verbs in parentheses.

Jack Taylor (enjoy) swimming, so he (go) to the beach last week.
He (stay) there for five days.
He (plan) to go back next year.

VI. Summary

Language tests support and give structure to any ESL program. They provide important insights into all aspects of a student's language proficiency, in all language skills, and they are an indispensible guide for placing a student in an appropriate class and instructional level. They help teachers plan the language curriculum and each class lesson, so that the particular needs of each student will be met. They are useful for students, too, because they set very specific goals, and they function as significant motivation for many language students. Finally, they are an aid in predicting whether a student is ready for the additional language demands of school, or job, or other training.

Yet it is important to remember that language tests test language, and they don't measure other factors, (like motivation), which are likely to have a powerful influence on successful achievement. An important decision regarding a student's future, such as whether the student is ready to enter a vocational training program, should not be made on the basis of a language test alone,

especially when it involves an area where the actual English language demands are not well understood. Teacher judgement should play a role as well. But language tests, used skillfully and appropriately, are important adjuncts to the design and functioning of an effective ESL program.

VII. Bibliography

The ESL tests described here are generally available and widely used in ESL programs for Indochinese students and other ESL students as well. Most come with comprehensive manuals that describe how the test was constructed and provide technical information regarding the test's validity and reliability. The manual will include detailed instructions about how to administer and score the tests, and they will indicate how to interpret the scores.

We have divided the tests into several categories: discrete-point tests; tests of oral fluency; and secure comprehensive tests. The first two terms are described in the Guide; the last refers to tests that are given several times a year, at designated locations, and are used primarily to test whether a student is prepared to enter a college or university program.

A. Discrete-Point Tests

Best, Jeanette, and Donna Ilyin. Structure Tests -- English Language (STEL). Rowley, Mass.: Newbury House, 1976. About $8.00 per packet. Answer keys, about $9.00 per set. Additional answer sheets, about $5.50 per set of 20.

Thirty-minute tests of English structure, which can be used for placement or as measure of general achievement. There are three sets of tests: Beginning I and II, Intermediate I and II, and Advanced I and II. A packet will consist of ten test booklets and ten answer sheets for forms I and II of a particular level, for a total of twenty test booklets and twenty answer sheets.

Brinson, Thomas C. Orientation in American English Placement Test. Silver Spring, MD: Institute for Modern Languages, no date. About $77. per packet; test specimen set, about $1.00.

Two-part placement test to determine level of ability in English. It was designed to place students in appropriate levels of the Orientation in American English text, but has been found to be a good placement guide whatever text is used. Students are tested orally, through questions and answers; and on comprehension, reading and writing through a written test. Oral test takes about ten to fifteen minutes per student; written test takes

about a half hour. Test packet consists of a teachers' guide, test booklet, 30 student test booklets, and 30 rating sheets.

Davidson, David M. Test of Ability to Subordinate. New York: Language Innovations, Inc., 1978. About $13.00 per packet. Additional answer sheets, about $2.25 per 100.

Test of ESL students' ability to combine sentences, for use as a diagnostic tool. Students are asked to fill blanks in sentences; test takes about half an hour to administer. Packet includes a teachers' manual, thirty test booklets, and sixty answer sheets.

Davis, Alva L. Diagnostic Test for Students of English as a Second Language. New York: McGraw-Hill, 1970. About $5.00 per packet of 10 test booklets; about $2.00 per packet of 10 answer sheets.

Forty-five minute test of 150 multiple choice items designed to diagnose areas of weakness in ESL students' command of English.

English Language Institute, University of Michigan. English Placement Test. Ann Arbor, MI: English Language Institute, University of Michigan, no date. Test packet, about $10.00. Specimen set available.

One hundred-item, multiple-choice test, intended for placement of students in beginning, intermediate, or advanced level classes. Test measures listening comprehension, grammar in conversational contexts, vocabulary and reading comprehension. Takes about 75 minutes to administer.

_____ Examination in Structure. Ann Arbor, MI: English Language Institute, University of Michigan, no date. Test packet, about $6.50. Specimen set available.

One hundred-fifty-item diagnostic test of knowledge of basic grammatical structures. Sixty-five per cent multiple choice, 35 per cent completion items. Forms A, B and C available; test takes about an hour. Packet consists of 20 test booklets, 100 answer sheets, and a two-part answer key.

_____ Michigan Test of Aural Comprehension. Ann Arbor, MI: English Language Institute, University of Michigan, no date. Test packet, about $9.00. Specimen set available.

Sixty-item test to measure understanding of spoken English. Three forms available; test takes about an hour to administer. One section requires students to choose pictures to match oral cues; with careful monitoring and supervision, this section can be used to test illiterates. Packet consists of a manual, 20 booklets, 100 answer sheets, and 3 scoring stencils.

_____ Michigan Test of English Language Proficiency (MTELP). Ann Arbor, MI: English Language Institute, University of Michigan, no date. Test packet, about $11.00. Specimen set available.

A three-part test of grammar, vocabulary and reading comprehension often required of university entrants. Can be used for placement, or as a general measure of achievement. Several forms are available, so the test can be used in before-and-after situations. One packet consists of one manual, 20 test booklets, 100 answer sheets, and an answer key. Test takes about an hour and a half to administer.

Harris, David P. and Leslie A. Palmer. Comprehensive English Language Test for Speakers of English as a Second Language (CELT). New York: McGraw-Hill, 1971.

Listening test with tapes, about $18.00
Structure test kit, about $10.00
Vocabulary test kit, about $10.00
Replacement test booklets, about $7.00 per packet of 20
Answer sheets, about $3.50 per packet of 100
Specimen sets for each test, about $3.00

Test of proficiency, especially appropriate for refugees with lots of education, and intermediate or advanced command of English. Listening test requires students to answer multiple-choice questions, and takes about 45 minutes. Vocabulary test has 75 multiple-choice questions, and requires about 30 minutes for administration. Can be used for placement, and as a general measure of achievement.

B. Tests of Oral Fluency

Ilyin, Donna. Ilyin Oral Interview. Rowley, MA: Newbury House, 1972. Test book about $14.50; score sheets, about $5.00 per packet of 50.

Test of students' oral comprehension and production through a series of questions geared to pictures. Questions become progressively harder, and test progressively more complex structures. Given to students individually, the test takes up to a half hour per student. Requires practice on the part of the examiner(s)!

Kunz, Linda. The John Test, A Test of Oral Proficiency for ESL Placement. New York: Language Innovations, Inc., 1976. About $3.50 per packet.

A quick placement test (named after the character in the test) widely used in refugee ESL programs, and especially appropriate for illiterate or little-educated refugees. Testing takes about five minutes per student. Packet includes 20 score sheets, a ditto master, pictures around which the questions center, and instructions.

C. Secure Tests

Educational Testing Service. SLEP, Secondary Level English Profiency Test for Non-native English Speakers. Administered by Educational Testing Services, Princeton, N.J.

SLEP is a test parallel to the TOEFL test annotated below, but designed for high school students. SLEP will be given, starting fall 1979, at particular centers in the United States, and security will be carefully controlled, as it is for the SAT's and other formal tests. For information, write SLEP Program Office, Room P235, ETS, Princeton, N.J., 08541.

_____ TOEFL, Test of English as a Foreign Language. Administered by Educational Testing Service, Princeton, N.J.

This is the famous TOEFL (pronounced toe-full, accent on first syllable) test, which is an entrance requirement for non-native speakers of English entering most American universities. Refugees with lots of education will run up against the TOEFL whenever they look for advanced training. The tests must be given at particular centers, as security is rigidly controlled. For information, write ETS, Box 899, Princeton, N.J., 08540.

Educational Testing Service, Test of Spoken English (TSE). Princeton, N.J.: Educational Testing Service.

A high-powered oral English test, from the same people who do the TOEFL. The test is currently in the validation stage, having been researched for the last two years. The test requires examinees to respond orally to a variety of printed and recorded stimuli, and takes about twenty minutes per student. For information, write to TSE Program Office, Room P229, Educational Testing Service, Princeton, N.J., 08541.

D. Tests in the Experimental Stage

Gonzales, Gustavo and Mary Galvan. Test of English for Adults of Limited English Speaking Ability. Available in 1980.

For use with adults enrolling in vocational training programs (particularly bilingual ones). Both a placement and achievement test.

Ilyin, Donna. Mini Tests. Rowley, MA: Newbury House. Available 1980.

Packets of short tests of specific units, e.g. grammar, vocabulary, telling time, etc. Covers beginning to advanced.

_____ Listening Comprehension Series. Rowley, MA: Newbury House. Available 1980.

A picture and a written test to measure listening comprehension. Beginning through advanced.

E. Additional Teacher Resources

For teachers who want to develop skills in test construction and validation, the following resources are useful. Language testing in a highly technical field, and constructing reliable and valid language tests is a demanding and time-consuming job. These resources are intended primarily to aid the teacher in the far less ambitious task of making well-designed tests and quizzes for the needs of the language classroom.

Harris, David P. Testing English as a Second Language. New York: McGraw-Hill, 1969. About $3.50.

Harris's book is a swift and comprehensive overview of the basic issues in ESL testing. It is mainly non-technical, and it has long been recognized as a standard introduction to language testing. It is now somewhat dated. It's not intended as a how-to book, but rather it presents examples of all aspects of the field in a concise way.

Valette, Rebecca M. Modern Language Testing. Second edition. New York: Harcourt Brace Jovanovitch, 1977.

This is a handbook on language testing for ESL teachers and foreign language teachers. It exhaustively catalogues various types of language tests and approaches to language testing. It really covers the field, with many examples. It is intended as an aid to the classroom teacher who needs to make tests for classroom use.

Bartz, Walter. Testing Oral Communication in the Foreign Language Classroom. Language in Education: Theory and Practice 17. Arlington, VA: Center for Applied Linguistics, 1979. $3.95.

This is a brief, up-to-date, and extremely useful guide to testing oral language fluency. Bartz defines the issues very clearly, and he offers a number of quite practical strategies for making and giving simple measures of student oral performance.